

Appendix A

Table 1: Biological analysis of 19 regulators:

For each regulator, gene name, number of regulatee genes, annotation and types of regulation (signaling, transcription) is detailed. Significant GO terms are listed as obtained from the automatic procedure. Verdict and comments are derived from manual inspection of the automatic annotation in light of the biomedical literature, and the specific regulates (visualized via GENESYS). PMIDs are used as references when needed.

Regulator name (# of regulates, signaling/TF)	Annotation (YPD, SGD)	Significant GO terms (significance, number of genes)	Verdict	Comments, selected regulatees
SST2 (99, signaling)	Regulator of G-protein signaling family member that negatively regulates the mating pheromone response pathway by binding to Gpa1p and stimulating its intrinsic GTPase activity, leading to desensitization to pheromone	Zygot formation (0,7) Cell-cell fusion (0,7) Mating (0,17) Genetic exchange (0,17) Developmental processes (0,17) Cell adhesion receptor (8.4e-05,3) Pheromone processing (8.4e-05,3) Signal transducer (8.9e-05,7) Transmembrane receptor (0.000119,4)	Proof of concept	Both the biological process (mating) and molecular function (signal transduction) is fully supported.
MET28 (254, TF)	Transcriptional activator regulating sulfur amino acid metabolism that functions with Met4p and Cbf1p, member of the basic leucine zipper (bZIP) family	Amino acid biosynthesis (0,16) Amino acid metabolism (0,24) Sulfur metabolism (1.2e-05,6) Sulfur utilization (1.2e-05,6) Nitrogen metabolism (1.5 e-05,7), ribosome biogenesis (0.000192,17), Oxireductase acting on sulfur group of donors/disulfide as acceptor (0.008938,2)	Proof of concept	Many methionine genes, <u>including additional transcription factors</u> , such as MET32.
GAT1 (125, TF)	GATA zinc finger transcription factor that plays a supplemental role to Gln3p, activating genes needed to use non-preferred nitrogen sources (nitrogen catabolite-repression sensitive genes)	Protein biosynthesis (0,30) Structural proteins of the ribosome (6e-06,18) Amino acid metabolism (0.000106,10) Hydrolase, acting on carbon-nitrogen (but not peptide) bonds in linear amides (0.005208,3) amine/polyamide transporters (0.006761,3) Nitrogen starvation response (0.008642,2) Amino acid transporter (0.010621,3) Other carbo-nitrogen ligase (0.024753,2)	Proof of concept	Both biological processes (nitrogen starvation response) and molecular mechanisms (protein biosynthesis, ribosome biogenesis, amino acid transport, carbo-nitrogen hydrolase and ligase) regulated by the gene are correctly predicted (PMID 11416207).

TEA1 (141, TF)	Ty1 enhancer activator of the Gal4p-type family of DNA-binding proteins	Amino acid and derivative metabolism (0,15) Amino acid biosynthesis (0,11) Amino acid metabolism (0,14) Nitrogen metabolism (7.9e-05,5) Urea cycle intermediate metabolism (0.000679,4) Carbon-nitrogen ligase\ with glutamine as amido-N-donor (0.001754,3)	Novel function assignment	Putative role: Regulation of amino acid biosynthesis (nitrogen utilization?). No past evidence suggests assignment to any biological process
UGA3 (73, TF)	Transcriptional activator for 4-aminobutyric acid (GABA) catabolic genes, including UGA4, UGA1, and UGS2	Amino acid biosynthesis (1e-06,8) Urea cycle intermediate metabolism (5.4e-05,4) Nitrogen metabolism (7.8e-05,4) L-amino acid transporter (0.001734,2) Transferase/transferring nitrogenous groups (0.002385,3) Amino acid transporter (0.002385,3)	Proof of concept	Regulated molecular mechanisms (amino acid transport, urea cycle, nitrogen metabolism) are correctly predicted. And distinguished from other GATA transcription factors (e.g. GAT1). Some specific target genes, such as the tryptophan permease TAT2, are specifically supported by previous evidence (PMID 10491387)
APG1 (168, signaling)	Serine/threonine protein kinase involved in induction of autophagy after nutrient limitation	Protein biosynthesis (0,29) Structural protein of ribosome (0,24) Hydrolase, hydrolyzing O-glycosyl compounds (0.000288,6) Carbon-nitrogen ligase, with glutamine as amido-N-donor (0.032127,2)	Proof of concept (partial)	The terms cover autophagy related and APG1-regulated events only partially. APG1 is induced by Tor proteins, which are known to regulate ribosome biosynthesis (PMID 11416184, 10995454). However, regulation by APG1 was not previously shown for this process. In addition to the terms, several regulatee genes that were previously associated with autophagy (AMS1, AUT2, AUT4, VPS34) provide further support.

SLT2 (245, signaling)	Serine/threonine protein kinase of MAP kinase family involved in the cell wall integrity (low-osmolarity) pathway.	Cell wall structural protein (4e-06,6) Cell wall organization and biogenesis (5.4e-05,11) Aspartic type endopeptidase (0.004971,3) Protein kinase cascade (0.017959,2) Mating (0.035059,8) Cell-cell fusion (0.038297,8)	Proof of concept	The biological process of cell wall modeling is supported by the first three GO terms. The genes in the latter three are consistent with SLT2's molecular function in signaling, including CMK2 and CMP2 from the calcineurin/calmodulin pathway which was shown to act synergistically with SLT2 (PMID 11230131, 7542741). Individual high-confidence genes further highlight SLT2's molecular function, and include RLM1, a TF known to be activated by SLT2 (PMID 10594829) and SSD1, a regulator of cell wall composition, and a multi-copy suppressor of SLT2 mutants (PMID 8662190, 11208019)
TPK1 (603, signaling)	Catalytic subunit of cAMP-dependent protein kinase 1, protein kinase A or PKA.	Ribosome biogenesis (7e-06,34) Protein biosynthesis (3.3e-05,58) Structural protein of ribosome (0.000128,46)	Proof of concept (partial)	The Ras/cAMP pathway was reported to induce the expression of ribosomal protein genes, in a predominantly transcriptional response (PMID 7760815). Other responses associated with the pathway were not observed.
SIP4 (265, TF)	Transcriptional activator of gluconeogenic genes through CSRE elements, activated by Snf1p kinase, contains a Zn[2]-Cys[6] fungal-type binuclear cluster domain.	Structural protein of ribosome (1.1e-05,28) Protein biosynthesis (6.9e-05,31)	Not supported	Only a handful of gluconeogenesis genes were found (GO term was not found to be significant). No

				previous support for regulation of ribosome biogenesis or protein synthesis by SIP4.
TEC1 (104, TF)	Transcriptional activator, involved with Ste12p in pseudohyphal formation	Mating (3.2e-05,9) Cell-cell fusion (3.7e-05,9) Developmental processes (9e-05,9) Pheromone response (9.2e-05,6) Cell surface receptor linked signal transduction (0.013253,3) Cell communication (0.028046,7) Actin binding (0.033924,2)	Partial support. Putative cross-talk.	The well-known role of TEC1's role in pseudohyphal growth is only marginally supported by individual regulatee genes (including STE12, its transcriptional counterpart, several cell cycle regulators, actin bundling proteins ARP3 and SAC6, and the heat shock TF HSF7). On the other hand, the prominent pheromone mating response, was not previously known to be regulated by TEC1. However, the two pathway cross-talk (they share a common TF, STE12, and the mating pathway was shown to activate TEC1, PMID 9657388). Thus an additional putative cross-talk point is identified.
KIN1 (292, signaling)	Serine/threonine kinase. Null mutant is viable and shows no obvious phenotype	Macromolecule biosynthesis (6.5e-05,36) Protein biosynthesis (0.000182,32) Structural protein of ribosome (0.001994,28) Cytokinesis (0.003699,6) Budding (0.005838,8) Axial budding (0.005925,5) Bud growth (0.007845,5) GTPase activator (0.019626,3) Cell cycle (0.034484,18)	Novel function assignment	Offers novel function assignment in cell cycle regulation (budding and cytokinesis) . Individual genes in these GO terms include multiple signaling and regulatory proteins involved in budding (e.g. CDC3, LAS17, PWP2, GIN4, ELM1, CDC42, RSRI, BEM2, MSB3,

				EGT2). There is no previous evidence on the biological process that KIN1 participates in.
GCN4 (156, TF)	Transcription factor of the basic leucine zipper (bZIP) family, regulates general control in response to amino acid or purine starvation	Chromosome segregation (0.00019,5) Nuclear division (0.000466,9) M-phase of mitotic cell cycle (0.000521,7) M-phase (0.001993,10) Shmooing (0.013181,2) Amino acid and derivative metabolism (0.014344,8) Macromolecule catabolism (0.015812,7) Cell cycle (0.017701,12)	No support. Putative alternative function?	Only few amino acid biosynthetic genes appear. On the other hand, several GO terms for M-phase processes (chromosome segregation, nuclear division and migration, shmooing) appear. We found no previous evidence for this regulatory role.
GAT2 (347,TF)	GATA zinc finger transcription factor	Ion transporter (0.000294,12) Cation transporter (0.000531,11) Monovalent inorganic cation transporter (0.000575,8) Carbon-oxygen lyase (0.000804,9) Energy derivation by oxidation of organic compounds (0.000824,15)	No support. Alternative function assignment ?	Classical GATA regulation was not observed. In addition to the ion transporters, most other GO terms included respiration genes. We found no previous evidence for the regulation of respiration by GATA type TFs.
XBP1 (439, TF)	Stress induced transcriptional repressor. Induced by stress and starvation during the mitotic cycle.	Heat shock protein (0.0028,9) Structural protein of ribosome (0.002512,32) Chaperone (0.003919,13) Sporulation (0.007804,6) Protein folding (0.009524,10) Stress response (0.014443,16)	Proof of concept. Extension of function assignment.	Both the biological process (stress response), and the regulated molecular function (chaperones, protein folding genes, heat shock proteins) were correctly predicted, including both effectors and other regulators (e.g. HAC1- TF for the unfolded protein response). The connection with sporulation (based mostly on starvation related genes) suggests an extension of XBP1 role, and is consistent

				with recent reports that XBP1 repression of CLN1 is important for efficient sporulation in yeast (PMID 10611226) as well as contributes to nitrogen starvation induced pseudohyphal growth (PMID 11340165).
YOL128C (170, signaling)	Protein kinase. Unknown biological process and molecular function. Homologous to GSK3 and MCK1 (PMID 10958699)	Cell communication (0.000362,14) Nitrogen starvation response (0.000944,3) Response to external stimulus (0.00426,8) Response to biotic stimulus (0.004271,6) Cell cycle control (0.020422,5) Small GTPase mediated signal transduction (0.031006,4) Cell cycle (0.031847,12) Signal transduction (0.034232,6)	Novel function assignment	Based on the GO terms, and the various cell cycle regulators they include, we postulate that YOL128C is involved in cell cycle regulation, perhaps in response to certain starvation signals . This is consistent with its homology to the cell cycle regulator GSK3.
DAL80 (196, TF)	GATA-type zinc finger transcriptional repressor for allantoin and 4-aminobutyric acid (GABA) catabolic genes. Null is deficient in allantoin degradation.	Hydrolase/ acting on carbon-nitrogen (but not peptide) bonds (0.000623,3) Transport (0.001813,22) Nitrogen starvation response (0.020239,2)	Proof of principle	Correct prediction of regulated biological process (nitrogen starvation response) and molecular targets (transporters, nitrogen catabolite regulated genes). Distinct from the other GATA factors, and specifically regulates the allantoin pathway genes (DAL 1,2,3, and 7), consistent with its proven role.. Regulates several transcription factors, including another GATA TF, GAT1, consistent with recent reports (PMIDs 11073899, 9171383).
IME4	Positive transcription factor for IME1 and	Chromosome segregation (0.001469,4)	Proof of	Role in cell cycle regulation

(145, TF)	IME2, mediates control of meiosis by carrying signals regarding mating type (a/alpha) and nutritional status.	Structural protein of cytoskeleton (0.006192,4) M phase of mitotic cell cycle (0.009698,5) DNA replication and chromosome cycle (0.018868,6) Nuclear migration (0.024512,2) Mitotic cell cycle (0.026572,7) M phase (0.036185,7) Cytokinesis (0.036317,3) Spindle pole body and microtubule cycle (0.04707,3)	concept (partial)	(in particular M-phase processes) is supported by dominant GO terms and their constituent genes (both regulators and structural genes). However, we do not observe meiosis specific genes, or IME4's direct target (IME1) which mediates this response.
MATALPHA2 (184, TF)	Homeobox-domain containing protein, which, in haploid cells, acts with MCM1 to repress a-specific genes. In diploid cells alpha2 acts together with a1 to repress transcription of haploid-specific genes.	Developmental processes (0.001573,10) Protein targeting (0.002185,9) Zygote formation (0.004601,4) Oxireductase\ acting on heme group of donors/oxygen as acceptor (0.011231,3) Mating type determination (0.017987,2) Membrane fusion (0.017987,2) Mating type specific transcriptional control (0.017987,2) Mating (0.023214,7) Cell-cell fusion (0.025269,7)	Proof of concept	The predominant GO terms are consistent with MATALPHA2's known function and refer to mating and related processes, including many mating-type specific genes a genes and other mating silencing counterparts (e.g. SIR2).
KIN82 (127, signaling)	Putative serine/threonine protein kinase most similar to cyclic nucleotide-dependent protein kinase subfamily and the protein kinase C subfamily. Biological process unknown.	Nucleotide metabolism (0.002014,3) Primary active transporter (0.003399,6) Mating-type specific transcriptional control (0.008909,2) Mating-type determination (0.008909,2)	Novel function assignment	No impressively predominant GO term. However, a recent screen (PMID 11337509) indicated that KIN82 may participate in pheromone response pathway , as its overexpression interferes with pheromone induced cell arrest. This is consistent with some of our GO terms.